

Durham Research Online

Deposited in DRO:

03 April 2020

Version of attached file:

Accepted Version

Peer-review status of attached file:

Peer-reviewed

Citation for published item:

Lopes, J. and Gomes, C. and Oliveira, C. and Elliott, J. (2020) 'Research studies on dyslexia : participant inclusion and exclusion criteria.', *European journal of special needs education.*, 35 (5). pp. 587-602.

Further information on publisher's website:

<https://doi.org/10.1080/08856257.2020.1732108>

Publisher's copyright statement:

This is an Accepted Manuscript of an article published by Taylor Francis in *European journal of special needs education* on 24 February 2020 available online: <http://www.tandfonline.com/10.1080/08856257.2020.1732108>

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full DRO policy](#) for further details.

Research studies on dyslexia: Participant inclusion and exclusion criteria

Dyslexia is a term widely used to describe reading characterized by problems with fluent and accurate letter or word recognition. Yet there is no consensus about the definition, origin, and diagnosis of dyslexia and the term is often used very differently by researchers and practitioners. In many cases, research findings are employed by clinicians in ways that are misleading and potentially counterproductive. The present study takes the form of an examination of participant samples included in studies of dyslexia ($n = 800$) over 20 years (2000-2019). The findings show that (1) researchers use a wide range of inclusion and exclusion criteria; that (2) IQ-reading achievement discrepancy is the most common inclusion criterion for dyslexia samples; (3) studies typically compare dyslexic samples to normal controls but not to other poor readers; (4) dyslexia seems to be employed as a catch-all term for poor readers in general, not as a term to define a specific type of poor reader. Finally, (5) dyslexia studies are very rarely published in educational journals.

Keywords: dyslexia; poor readers; sample constitution; inclusionary criteria; exclusionary criteria.

Research studies on dyslexia: Participant inclusion and exclusion criteria

Since the term dyslexia was first coined by the German ophthalmologist. Rudolf Berlin, in 1887, many terms have been used to describe the difficulties that a significant proportion of people encounter in trying to decode text. Early terms included congenital word blindness (Hinshelwood 1907), visual agnosia for words, and word blindness (Orton 1925). More recently, terms commonly used include: specific reading difficulties, specific learning difficulties, learning disability, specific reading retardation, unexpected reading difficulty, reading disorder, reading disability, and specific reading disability.

Dyslexia has become one of the most widely employed terms to describe reading (decoding) and spelling problems. While this term has long been prevalent in the UK, its use has grown rapidly across the world, particularly in the United States where dyslexia advocacy and legislation have flourished (Gabriel 2019; Ward-Lonergan and Duthie 2018). Despite its ubiquity, achieving a workable and scientifically valid definition proved difficult (Elliott and Grigorenko 2014a). What has compounded the difficulty is that there is often a gulf between understandings of the construct of dyslexia in the research literature and formal definitions put forward by practitioner and advocacy groups. This has impacted upon the ability of professionals to operationalize the term for the purposes of educational, clinical and even forensic practice (Knight 2018; Ryder and Norwich 2018).

For many researchers, the term dyslexia should be seen as synonymous with reading disability (i.e. a severe difficulty in decoding text) (Elliott and Grigorenko 2014a, 2014b; Protopapas and Parrila 2018; Snowling et al. 2019). Thus, Peterson and Pennington (2015) state:

“Dyslexia is mainly defined as the low end of a normal distribution of word reading ability....Thus, in order to diagnose the disorder, a somewhat arbitrary cut-off must be set on a continuous variable” (p. 285).

Similarly, Seidenberg (2017) states:

“Dyslexics are children (and later adults) whose reading is at the low end of a normal distribution. Reading skill results from a combination of dimensional factors (that is, ones that vary in degree), yielding a bell-shaped curve. The reading difficulties of the children in the lower tail are severe and require special attention. ‘Dyslexia’ refers to these children. Viewed this way, dyslexia is on a continuum with normal ‘reading’. All children face the same challenges in learning to read, but dyslexics have more difficulty with the essential components” (pp.156-157).

Within the United States, the notion that dyslexia is manifested by an ‘unexpected’ problem with reading has been widely put forward as a means to differentiate between dyslexic and other poor readers (Stanovich 2005; Elliott and Nicholson 2016; Vellutino et al. 2004). Here, it is suggested that reading problems may be expected for some children (the poor readers, or the “garden-variety poor readers” in the words of Stanovich (1988), but not for others (the dyslexics) (Christo 2015; Ferrer et al. 2010). The notion of unexpectedness is at the origin of the IQ-discrepancy model of specific learning disabilities (Beaujean et al. 2018). According to this perspective, reading problems are unexpected when the individual has a normal or above normal IQ. Shaywitz (2003) suggests that unexpectedness is revealed by an uneven cognitive profile in which a decoding weakness is typically surrounded by a 'sea of strengths' (p. 58), not usually found in other poor readers. These may include high functioning with

respects to reasoning, problem-solving, critical thinking, vocabulary, comprehension and general knowledge.

While such an account might appeal to those with such strengths, there is little evidence that the reading difficulties of others with flatter profiles are inherently different in respect of appropriate intervention or prognosis (Fletcher et al. 2013). Furthermore, socially disadvantaged children with reading disabilities are likely to experience particularly acute problems with general knowledge, comprehension, and vocabulary, often attributable to reduced access to the life opportunities that could compensate for their lack of access to the written word (Kodan and Akyol 2018; Lee and Kim 2015). Reduced vocabulary will hinder the child's ability to decode unknown printed words, particularly where these are partially decoded or irregularly spelt and, as a result, further hinder the development of the child's phonological recoding skills (Lawrence et al. 2018; Tunmer and Greaney 2010).

Given the flaws of the IQ discrepancy model (Fletcher et al. 2007; Gresham and Vellutino 2010; Stuebing et al. 2009), it is unsurprising that this model is now generally agreed to be inappropriate for diagnosing dyslexia. Thus, the 5th edition of the Diagnostic and Statistical Manual (DSM-V), produced by the American Psychiatric Association (2013), dispensed with such an approach for identifying learning disabilities in reading and other academic areas. Similarly, major dyslexia advocacy groups, such as the International Dyslexia Association and the British Dyslexia Association, contend that dyslexia features across the intellectual spectrum. Interestingly, however, despite this, the discrepancy model continues to be used by some practitioners and researchers (Machek and Nelson 2007; O'Donnell and Miller 2011).

An alternative way to identify an unexpected problem with reading is to examine the intractability of the condition for a given individual following provision with high-quality, evidence-based intervention over time (Fletcher et al. 2018). By definition, such a determination could not take place at an early age, or before several years of intense intervention had taken place. While this conception is increasingly popular for determining eligibility for special education in the United States, it is very rarely employed as a means of selecting participants for dyslexia research studies.

A particular difficulty exists where dyslexia researchers draw upon a pool of poor readers as participants for their studies (e.g. in cognitive science, neuroscience, genetics, or education), but apply subsequent findings solely to a dyslexic subgroup. For example, if, as is often the case, one were to recruit those reading at or below the 15th centile on a standardised reading test for a study of various cognitive processes (e.g. working memory), it would be invidious to subsequently claim that, should the observed mean performance be below that of a comparator group of normal readers, such processes could be considered to be markers of dyslexia, and thus these could then be employed to differentiate dyslexic individuals from other poor readers. However, such circular-definition practices are common (Elliott and Grigorenko 2014a).

Constitution of participant samples in studies of dyslexia

It is often unclear in dyslexia studies whether the use of this term applies to all struggling decoders, or to a subset that can be differentiated from other poor readers. Where there is a suggestion that the latter case applies, criteria for making this distinction are rarely made explicit (Lopes et al. 2016). Nergård-Nilssen and Eklund (2018), for example, recently developed an instrument, “*The Norwegian screening test for dyslexia*”, that was designed to detect dyslexia in upper secondary school students.

Participants in the development of the instrument were divided into non-impaired and impaired reading groups through a self-report questionnaire. Of the 38 ‘impaired’ students, twelve were diagnosed with dyslexia. However, the authors do not report who diagnosed the students and simply note that: “Unfortunately, we do not have access to their test protocols or test reports” (p. 4). Thus, the reader cannot scrutinise how participants diagnosed with dyslexia actually differed from other participants who self-reported reading problems. Furthermore, results were presented for the ‘impaired’ and for the ‘non-impaired’ group, but not for the subsequently identified specific group of dyslexics.

Reviews of studies with adults have shown that highly diverse criteria are often employed (e.g., Rice and Brooks 2004). A similar difficulty seems to apply in studies conducted with children and youth. For instance, several studies with dyslexic individuals report that participants have a diagnosis of dyslexia but offer no further specifications (Stagg, Eaton, and Sjoblom 2018; Sjoblom, Eaton, and Stagg 2016), while other studies essentially detect dyslexic students by their low performance on one or more reading tests (Diamanti et al. 2018; Breadmore and Carroll 2016). In other cases, the terms “dyslexics” and “poor readers” seem to be taken as equivalents or are taken as synonymous, but, somewhat puzzlingly, some differentiation between dyslexia and poor reading is still suggested (e.g., Uno et al. 2009).

Other issues might affect the composition of samples of dyslexic subjects. For instance, even where studies present clear criteria, these are highly variable across different studies. Moreover, even for a specific study, children included in the same group, on the basis of the same selection criteria, may present considerable phenotypic variations. Within-disorder variations, however, are discarded when groups are statistically compared (Peters and Ansari 2019).

More than three decades ago, Ellis and Large (1987) considered that bringing together children of different ages, with different backgrounds and taught in different ways, would not enable the identification of a putative average reader. These authors stated that either the group is completely homogeneous (which hardly happens), or the individual pattern is obscured by the group average. The group average does not represent the performance of the individual and does not permit group generalizations for the individual. Lyon and Moats (1997) also highlighted the limitations and challenges of sampling strategies and study designs in reading intervention research. These included sample heterogeneity and vague inclusion criteria, poorly defined interventions, inadequate control groups, inadequate intervention and transfer effects, poorly controlled teacher effects, low fidelity of program implementation by teachers, and inadequate measurement practices.

Much has been done in the last decades in the area of reading research and dyslexia and the design of studies has become progressively more sophisticated (see, for instance, Aravena et al. 2016; Tilanus, Segers, and Verhoeven 2019a, 2019b). Approaches that have sought single factor explanations of reading disability/dyslexia fail to address the complexities involved and it is now understood that multiple deficit model approach is required (e.g., Pennington 2006; Kauffman 2004; Ring and Black 2018). In this respect, "... a multiple-deficit model proposes that the development of any complex cognitive behaviour, such as reading, is the outcome of multiple levels of interacting factors that may be protective or increase risk of a disorder" (Ring and Black, p. 2). That approach must acknowledge that the etiology of the problem is multifactorial, that comorbidity is the rule, not the exception, and that the distribution of the disorder is often continuous, not discrete (therefore, cut-off points are usually arbitrary). Moreover, longitudinal and training studies, are important to map the course

of the disorder and to identify the factors that can mediate or moderate its development (Lyon and Moats 1997; Diamanti et al. 2018; Snowling et al. 2019; Tong et al. 2017).

Ensuring the appropriate selection of participants for studies of reading disability presents a number of methodological challenges and it is important that researchers recognize and acknowledge this.

The present study

It appears that the term “dyslexia” is understood very differently by researchers, which leads to a lack of consistency and clarity in the make-up of participant groupings for dyslexia research studies. While this difficulty has been shown for adult studies, albeit more than fifteen years ago (Rice & Brooks, 2004), the picture for children and young people is unclear. Accordingly, the main goal of the present study was to identify the nature of participants who have been included in studies of dyslexia since 2000. Specifically, we sought:

- (a) To identify selection criteria in studies of dyslexia, and to assess the degree of homogeneity of the criteria used for the selection of samples;
- (b) To compare the selection criteria for samples of struggling readers (other than ‘dyslexics’) with selection criteria for samples of identified dyslexics;
- (c) To identify the contexts from within which samples of dyslexics were collected;
- (d) To identify the professional characteristics of groups involved in the diagnosis of participants for studies on dyslexia.
- (e) To review the characteristics of the scientific journals that publish studies on dyslexia.

Method

Sample

Eight hundred journal articles involving dyslexia were analyzed. In 764 of these, we found samples of either ‘dyslexics only’ or samples that contrasted ‘dyslexics’ and ‘normal readers’, or some other subjects (e.g., speech and language impaired). Seven studies included samples of ‘dyslexics’ that were contrasted with samples of other struggling readers and with normal controls.

Procedure

An analytical grid was developed to examine the following features of each journal article:

- a) Title of the article;
- b) Article publication year;
- c) Journal in which the article was published;
- d) Number of identified dyslexic participants in the sample;
- e) Contexts of sample collection;
- f) Criteria for sample selection;
- g) Test battery for sample selection;
- h) The professional categories of groups involved in the diagnosis of participants for studies on dyslexia.

Studies involving samples of identified dyslexic participants were retrieved from the Elsevier SCOPUS database of more than 16,500 peer-reviewed journals. Journal articles were searched using the keywords Dyslexia or Dyslexics. The search was limited to Article in the Document Type field. From the whole set of articles retrieved from SCOPUS, the 40 most cited articles each year from 2000 through 2019 were retained for additional analysis.

We analyzed empirical studies in the database that involved (a) samples of identified dyslexic participants and samples of normal readers, (b) samples including both identified dyslexic participants and struggling readers not identified as dyslexic, and (c) samples of dyslexic readers only. Studies (a) of struggling readers who were not held to be dyslexic and (b) non-empirical studies (e.g. systematic reviews, meta-analyses, editorials, and commentaries) were excluded.

Data analysis

Both qualitative and quantitative methods were used for data analysis. Qualitative content analysis was used initially (Cho and Lee 2014) to elicit general categories, and this was followed by categorical content analysis. According to Patton (2015), this kind of analysis makes use of objective and systematic procedures that allow the researcher to make logical deductions from the selected data corpus.

Qualitative analysis was followed by quantitative analysis of the data whereby the frequency of each of the identified categories was calculated. Low frequency (below 5%) and/or a lack of conceptual relevance were identified as criteria for eliminating some categories. Only one study was excluded on this basis.

For each of the final categories, the percentage of inter-observer agreement (two judges were used) was calculated by dividing the number of agreements by the total number of answers. Cohen's Kappa was not used because it was considered highly unlikely that raters would provide random answers given that the rating system referred to the presence or absence of a specific category (e.g., medical sciences, neurosciences, psychology, education). Rater disagreements were discussed and reviewed by both raters. When an agreement was not subsequently achieved, a third rater was called upon to assist. For an initial exercise involving 200 studies, the inter-rater agreement varied between

94% and 97% with a mean agreement of 95.8%. For this reason, it was decided that further analysis of this kind was not required.

Results

Criteria for the selection of samples

Inclusionary criteria: Table 1 shows the frequencies of inclusion criteria used by researchers for sample selection.

Table 1

Inclusionary criteria in sample selection: frequency and type (N = 800 studies)

	<u>Criteria</u>				
	Formal diagnosis of <u>dyslexia</u>	Results in a battery of <u>tests</u> ^c	School records of <u>learning problems</u>	<u>IQ</u>	No <u>information</u>
Dyslexic (800 ^a -100%)	315 ^b (39%)	596 (75%)	20 (3%)	498 (62%)	18 (2%)
Normal readers (738 - 92%)	-	487 (66%)	14 (2%)	370 (50%)	76 (10%)
Struggling readers (7 – 0.9%)	-	4 (57%)	-	5 (71%)	0

^a Number of studies including dyslexic readers; ^b Number of studies in which ‘dyslexics’ receive a formal diagnosis;

Note: categories are not mutually exclusive. A particular participant within the dyslexic group, for instance, might have received a formal diagnosis of dyslexia, might have undertaken several tests, and might have a school record of learning problems.

What becomes most apparent from Table 1 is that the criteria used to select dyslexic participants are stricter than the criteria used to select normal readers (the widely used control group). Also, there is only a very small number of studies with samples of poor readers other than people diagnosed as having dyslexia.

The battery of tests employed typically included cognitive (usually IQ) and reading tests. The WPPSI-R, the WISC-R, the WAIS and the Raven scales are by far the most commonly used cognitive measures. The Stanford-Binet, the Woodcock-Johnson Test of

Cognitive Abilities, The Primary Test of Cognitive Skills, or the KBIT Test (Kaufman Brief Intelligence Test) were less frequently used measures in the 800 studies that were reviewed. Reading tests comprised a range of standardized measures such as the Woodcock-Johnson Educational Battery, the Reading and Spelling subtests of the Wide Range Achievement Test, the Alouette's standardized reading test, or the Hong-Kong Test of Reading and Writing (HKY-SpLD), in addition to various non-standardized measures.

Exclusionary criteria: Table 2 shows the frequencies of exclusionary criteria used by researchers for sample selection.

Table 2

Exclusionary criteria in sample selection: frequency and type (N = 800 studies)

	<u>Criteria</u>					
	<u>Sensory deficits</u>	<u>Comorbidity</u>	<u>Other problems</u>	<u>Lack of educational opportunities</u>	<u>IQ</u>	<u>Not first language</u>
Dyslexic (800 – 100%)	226 (28%)	283 (35%)	78 (10%)	62 (8%)	33 (41%)	122 (15%)
Normal readers (738 - 92%)	162 (22%)	192 (26%)	60 (8%)	42 (6%)	28 (4%)	116 (16%)
Struggling readers (7 – 0.8%)	3 (43%)	7 (100%)	-	1 (14%)	2 (29%)	4 (57%)

Note: categories are not mutually exclusive.

Sensory deficits (hearing, vision), and comorbidities (e.g., ADHD, emotional disorders, mental illness, neurological disorders), are by far the most frequent exclusionary criteria. However, in some studies, conditions such as ADHD (attention-deficit hyperactivity disorder), are used as inclusionary criteria because the researchers intended to study subjects with comorbid dyslexia and ADHD. A significant number of studies refer that the language employed in the education system might not be an individual's first language and, in such cases, this is an exclusionary criterion. The language of the great majority of the samples is English (most frequently in UK or US settings) although there are also many studies in Finnish, Norwegian, Chinese, French,

Hebrew, German, Dutch, and Italian, and only a few in Spanish, Portuguese or Arabic, amongst others.

Contexts

Table 3 shows the frequencies of the contexts where samples were collected.

Table 3

Frequencies of the contexts in sample selection (N = 800 studies)

	<u>Clinical</u>	<u>Community</u>	<u>Re-used</u>	<u>No information</u>
Dyslexic (800 – 100%)	252 (28%)	391 (28%)	77 (28%)	112 (28%)
Normal readers (738 - 92%)	28 (4%)	445 (60%)	64 (9%)	146 (20%)
Struggling readers (7 – 0.8%)	2 (29%)	6 (86%)	-	1 (14%)

Note: Categories are not mutually exclusive.

The most noticeable fact in Table 3 is that a large number of studies provide no information about the locations from where samples were collected. It is also important to note that, in 28 studies, normal readers were recruited in clinical settings.

Professionals

Table 4 shows the frequencies of the professional groups that select/diagnose participants for studies on dyslexia.

It is important to note that in almost three-quarters of the studies, the researchers took responsibility for identifying whether participants were deemed to be dyslexic.

Research methodology

The results show that 98% (n = 744) of the studies represent basic research (e.g., genetic studies, spatial and temporal attention in developmental dyslexia,

Table 4

Professional groups that provide the classification/diagnosis

	<u>Researcher</u>	Unspecified professional in clinical or community context ^a	<u>Psychologist</u>	<u>Others</u>	<u>No information</u>
Dyslexic (800 – 100%)	529 (66%)	184 (23%)	65 (8%)	236 (30%)	45 (6%)
Normal readers (738 - 92%)	457 (62%)	133 (18%)	4 (0.5%)	42 (6%)	43 (6%)
Struggling readers (7 – 0.8%)	6 (86%)		-		1 (14%)

Note: Categories are not mutually exclusive

neuroanatomical studies) and that only 2% (n = 8) represent applied research, aimed to address pedagogical issues (e.g., classroom paired learning, music training to improve phonological awareness and reading skills, sublexical training with rhythmic background).

Journals that publish studies on dyslexia

The data provide evidence that empirical studies of dyslexia are most commonly found in neuroscience journals (43%), followed by psychology journals (24%), medical

sciences journals (14%), interdisciplinary journals (19%) and education journals (6%). Taken together, journals in the domains of neuroscience (a domain close to the medical sciences) and medical sciences publish more than half the studies. Perhaps the most striking finding is that education journals published a mere 5% of the studies in this field.

Sample size in studies on dyslexia

The results show that studies in this area typically used small samples. Almost 50% of the studies (n = 398); contained 1-50 participants; 25% (n = 202) had 50-100 participants; 21% (n = 165) had more than 100 participants; and 3% (n = 22) had more than 1000 participants. It should be noted that, in these samples, only about half of the participants are identified as dyslexic. The remainder are normal readers or, in seven of the eight hundred studies, perceived as some of kind of poor reader. Given that dyslexia is considered by some to affect as many as 20% of the population (Shaywitz 2003), the use of small sample sizes cannot be a function of the prevalence of reading difficulties. Sixty-two studies (8%) used only one group of participants (dyslexics), 529 (66%) used two groups (usually dyslexics + normal controls), and 175 (22%) used three groups (in most cases, the third group includes speech and language impaired children). Only seven (less than 1%) studies included poor readers and identified dyslexics.

Discussion

Inclusionary and Exclusionary Criteria in Studies of Dyslexia

Inclusionary criteria

Our results indicate that the inclusion of participants in studies of dyslexia typically relies upon performance on a battery of cognitive and/or reading tests. IQ-

Performance discrepancy was the most frequently used criterion in the 800 reviewed studies, despite recognition that such a criterion is currently inadequate (e.g., Cotton, Crewther, and Crewther 2005; Sternberg and Grigorenko 2002; Vellutino et al. 2004; Beaujean et al. 2018). As Schneider and Kaufman (2017) note, “Defining learning disability as unexpectedly low academic achievement despite average or better IQ has intuitive appeal, but it focuses on distal and relatively intractable influences on academic ability” (p. 3). Still, according to our review, in 2000, 24 studies out of 40 used the IQ-Performance discrepancy as an inclusionary criterion, compared to only nine in 2017, 13 in 2018, and 11 studies in 2019 (but 32 in 2016). Although these numbers suggest that something might be changing in the selection of samples of dyslexics, and perhaps in the clinical practice, it is too early to draw definite conclusions.

Other criteria found in many studies to select dyslexic participants include the formal diagnosis of dyslexia and a personal history of reading problems. Nevertheless, in a quarter of the studies, inclusionary criteria are unclear and quite variable. Sometimes it is even hard to understand the nature and the number of the participants. This lack of clarity conflicts with the expectation that inclusion criteria are explicit, extensively applied, and in line with current scientific knowledge (Scruggs and Mastropieri 2002).

Exclusionary criteria

Our review shows that many studies report a variety of exclusionary criteria, although some authors consider this practice as inadequate (e.g., Rutter and Maughan 2005; Fletcher 2009). Interestingly in most of the cases we examined, the criteria were

reported as characteristics of the sample (e.g. normal hearing, right-handed, no emotional problems).

Although sensory deficits and comorbidities are by far the most frequent exclusionary criteria, the lack of educational opportunities/poor schooling is an interesting and controversial issue. As Elliott and Grigorenko (2014b) point out, deciding whether educational experience can be largely blamed for an individual's reading difficulties is generally impossible, at least in the developed world. Indeed there seems to be no reliable way to assure that a student had enough educational opportunities beyond the finding that he or she attended school. Perhaps this explains why, despite widespread reference to this criterion, explicit exclusion from studies on these grounds is rarely, if ever, reported (Lopes et al. 2016). In the studies that we reviewed no specific evidence was presented about how the appropriateness of reading instruction was measured, or about how educational opportunities were conceptualized or approached. The results also show that IQ was used by many researchers both as an exclusionary (as well as an inclusionary) criterion.

Selection criteria for samples of struggling readers other than dyslexics

The results show that researchers typically use more criteria to select dyslexic individuals (including a formal diagnosis, not found for other poor readers) than to select other poor readers. However, Shaywitz and Shaywitz (2013) hold that the specific features that may lead to a diagnosis of dyslexia or to a diagnosis of any other poor reading condition are not clear. This issue is critical for the validity of research studies of dyslexia, and its use as a clinical category, unless, of course, this term is being used as a synonym for poor reading. "To test the hypotheses that developmental dyslexics differ from both normal readers and 'ordinary' poor readers, researchers need two

control groups, not one” (Rice and Brooks 2004, p. 33). Our review shows that this is almost never done. We found only 7 out of 800 studies in which there were two control groups: a group of normal controls and a group of poor readers other than dyslexics. In most of these seven studies, dyslexics and other poor readers differ in IQ scores but are matched by reading performance, a practice that is now considered inappropriate (Fletcher et al. 2007; Gresham and Vellutino 2010; Stuebing et al. 2009). Moreover, in the 800 studies, dyslexics are often interchangeably labelled as poor readers, delayed readers or struggling readers, thus suggesting that many researchers use dyslexia as a catch-all term for poor readers in general, not as a term to define a specific type of readers.

Contexts

Clinical and community settings are the most usual contexts for participants’ recruitment for studies of dyslexia. However, in almost two-thirds of the reviewed studies, the researchers themselves took charge of the selection process.

Interestingly we found no samples of poor readers (as contrasted with dyslexic individuals) selected in clinical contexts, with the resultant unfortunate bifurcation of struggling readers into those with a clinical v non-clinical condition, and the implications this has for the provision of intervention services. Authors such as Fletcher et al. (2007) have long noted that dyslexia, often presented as a medical condition, led to a schism between the medical and the educational communities and almost excluded the educational community from dyslexia/poor reading debate.

Professionals

Although our results show that the researchers themselves most commonly conduct the selection of samples of dyslexics, a significant number of other

professionals are involved in the selection/diagnosis process. Interestingly, teachers are very rarely considered as potential diagnosticians. At best, they are called upon to indicate which of their students show problems with printed words. This seems logical if one accepts that dyslexia is a neurological condition, but it is questionable since teachers are the most experienced professionals dealing with reading problems, and because effective interventions in reading problems are almost exclusively pedagogical (McGuinness 2005).

Journals that publish studies on dyslexia

The results show that almost 90% of the studies of dyslexia are published in medical, neuropsychological or psychological journals, with only 6% published in educational journals. More than 40% ($n = 342$) were published in neuroscience journals. In a previous study, Lopes (2012) found that “dyslexia research is overwhelmingly published in medical or psychologically oriented journals, with a strong neuropsychological perspective” and that “As important as, or even more important than, the number of published articles is the impact of each of these journals” (p. 3).

The finding that only 6% of the reviewed studies are published in educational journals is rather puzzling. Even considering that in most studies, dyslexia is conceptualized as “neurological in origin...” (e.g., Norton, Beach, and Gabrieli 2015; Richlan 2014; Xia et al. 2016; Su et al. 2018; Luciano et al. 2018; Di Liberto et al. 2018) we could expect a more balanced number of publications between medical and educational articles. It is important to take note of this imbalance because it seems to reflect a more general imbalance of the concept of dyslexia itself. Specifically, dyslexia is conceptualized as a condition that (a) is neurological in origin (i.e. somehow a disease), (b) specifically affects reading (that is, depends on a social

interaction/condition), (c) is essentially diagnosed through a reading test (any other tests are supplementary and ultimately expendable), and (d) requires instruction above any other intervention strategy. That is, dyslexia is mostly conceptualized as a medical condition that is better treated by instructional methods.

The relative absence of papers in educational journals is potentially problematic if research that addresses practitioner concerns about effective teaching methods is given lower consideration and status. It is now more widely understood that interventions for complex reading difficulties need to be based upon educational practices rather than on the basis of psychological, neuropsychological, or medical approaches (Fletcher and Wagner 2014; Moats 2014).

Conclusion

In conclusion, the picture that emerges is one of significant inconsistency in the selection of participants for studies of dyslexia. In most cases, researchers seem to use the term dyslexia to describe anyone who is struggling with reading (and sometimes spelling). Interestingly, our findings of dyslexia samples seem to parallel the problems that have been found in the more general field of learning disabilities (LD). In a review of studies published in the period 2001-2013, involving samples of students with LD, Williams et al. (2016) found that “nearly one-quarter of all identified studies investigating LD did not describe who identified the participants as having LD or how they were identified” (p. 226), a proportion that is similar to what was found by Durrant (1994) more than two decades earlier. These authors also found a generalized lack of control over identification procedures, as well as persistence in the use of the IQ-Achievement discrepancy method (although the proportion of studies using this method “drops precipitously” after 2010). Williams et al. (2016) therefore conclude that it is essential that researchers “move toward more complete descriptions of participants with

LD, including information about the criteria with which they were identified and the persons by whom they were identified”. (p. 228).

Our results indicate that explicit criteria for the composition of research samples of designated dyslexics are many times uncertain, or are highly questionable. In many cases, it is unclear whether the authors are using the term dyslexia as a synonym for poor reading or to describe a subset of poor readers. Where they do seek to make such a distinction, it is incumbent upon them to provide clear information about the specificity of their dyslexic participants in comparison with other struggling readers. Such detail, however, may well lead to challenge on the grounds that their criteria are invalid and inappropriate (Elliott and Grigorenko 2014a; Elliott and Nicholson 2016). It is important to note that clinicians often use findings from dyslexia studies to justify their diagnostic differentiations between dyslexic and non-dyslexic poor readers. However, this practice makes little sense if the findings that they are citing when doing this are related not to a dyslexic subgroup, but to the full heterogeneous group of struggling readers.

Limitations and Future Studies

This study presents a main limitation: only a small number of studies (about 6%) identify dyslexic participants who are then contrasted with other struggling readers. While this necessarily limits our ability to compare the criteria used in participant selection, it also points to the fragile research basis for making such a distinction. This limitation suggest that future studies should deeply explore the few studies that include dyslexic subjects and other poor readers or studies with other poor readers only to clearly understand what are the inclusion criteria in any of these groups. The specificity of dyslexia hardly hold until unique inclusion criteria are se

References

- American Psychiatric Association (APA). (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.
- Aravena, S., J. Tijms, P. Snellings, and M. W. van der Molen. 2016. "Predicting responsiveness to intervention in dyslexia using dynamic assessment." *Learning and Individual Differences* 49:209-15. doi: 10.1016/j.lindif.2016.06.024.
- Beaujean, A. Alexander, Nicholas F. Benson, Ryan J. McGill, and Stefan C. Dombrowski. 2018. "A Misuse of IQ Scores: Using the Dual Discrepancy/Consistency Model for Identifying Specific Learning Disabilities." *Journal of Intelligence* 6:1-25. doi: 10.3390/jintelligence6030036.
- Breadmore, H. L., and J. M. Carroll. 2016. "Morphological spelling in spite of phonological deficits: Evidence from children with dyslexia and otitis media." *Applied Psycholinguistics* 37 (6):1439-60. doi: 10.1017/S0142716416000072.
- Cho, J. Y., and E. H. Lee. 2014. "Reducing confusion about grounded theory and qualitative content analysis: Similarities and differences." *Qualitative Report* 19 (32):1-20.
- Christo, C. 2015. "Developmental dyslexia." In *The Cambridge Handbook of Communication Disorders*, edited by L. Cummings, 88-108. Cambridge: Cambridge University Press.
- Cotton, S. M., D. P. Crewther, and S. G. Crewther. 2005. "Measurement error: Implications for diagnosis and discrepancy models of developmental dyslexia." *Dyslexia* 11:186-202.
- Di Liberto, G. M., V. Peter, M. Kalashnikova, U. Goswami, D. Burnham, and E. C. Lalor. 2018. "Atypical cortical entrainment to speech in the right hemisphere underpins phonemic deficits in dyslexia." *NeuroImage* 175:70-9. doi: 10.1016/j.neuroimage.2018.03.072.
- Diamanti, V., N. Goulondris, M. Stuart, R. Campbell, and A. Protopapas. 2018. "Tracking the effects of dyslexia in reading and spelling development: A longitudinal study of Greek readers." *Dyslexia* 24 (2):170-89. doi: 10.1002/dys.1578.
- Durrant, J. E. 1994. "A decade of research on learning disabilities: A report on the state of the literature." *Journal of Learning Disabilities* 27 (1):25-33.
- Elliott, J. G., and E. L. Grigorenko. 2014a. *The Dyslexia debate*. New York: Cambridge University Press.
- . 2014b. "The end of dyslexia?" *Psychologist* 27 (8):576-80.
- Elliott, J. G., and R. Nicholson. 2016. *Dyslexia: Developing the debate*. London: Bloomsbury Press.
- Ellis, N., and B. Large. 1987. "The development of reading: As you seek so shall you find." *British Journal of Psychology* 78 (1):1-28. doi: 10.1111/j.2044-8295.1987.tb02222.x.
- Ferrer, E., B. A. Shaywitz, J. M. Holahan, K. Marchione, and S. E. Shaywitz. 2010. "Uncoupling of reading and iq over time: Empirical evidence for a definition of dyslexia." *Psychological Science* 21 (1):93-101. doi: 10.1177/0956797609354084.
- Fletcher, J. M. 2009. "Dyslexia: The evolution of a scientific concept." *Journal of the International Neuropsychological Society* 15 (4):501-8. doi: 10.1017/S1355617709090900.
- Fletcher, J. M., G. R. Lyon, L. S. Fuchs, and M. A. Barnes. 2007. *Learning Disabilities: From identification to intervention*. New York: Guilford Press.

- Fletcher, J. M., G. R. Lyon, L. S. Fuchs, and M. A. Barnes. 2018. *Learning disabilities: From identification to intervention*. New York: Guilford.
- Fletcher, J. M., K. K. Stuebing, R. D. Morris, and G. R. Lyon. 2013. "Classification and definition of learning disabilities: A hybrid model." In *Handbook of learning disabilities* edited by H. L. Swanson, K. R. Harris and S. Graham, 33-50. New York: Guilford Press.
- Fletcher, J. M., and R. K. Wagner. 2014. "Accumulating knowledge: When are reading intervention results meaningful?" *Journal of Research on Educational Effectiveness* 7 (3):294-9. doi: 10.1080/19345747.2014.925310.
- Gabriel, R. E. 2019. "Converting to Privatization: A Discourse Analysis of Dyslexia Policy Narratives." *American Educational Research Journal*. doi: 10.3102/0002831219861945.
- Gresham, F. M., and F. R. Vellutino. 2010. "What is the role of intelligence in the identification of specific learning disabilities? Issues and clarifications." *Learning Disabilities Research & Practice* 25 (4):194-206.
- Hinshelwood, J. 1907. "Four cases of congenital word blindness." *British Medical Journal* 2:1229-32.
- Kauffman, J. M. . 2004. "Introduction." In *Handbook of Research in Emotional and Behavioral Disorders*, edited by B. Rutherford Robert, M. Quinn Mary, Sarup R. Mathur and 11-31. New York: The Guilford Press.
- Knight, C. 2018. "What is dyslexia? An exploration of the relationship between teachers' understandings of dyslexia and their training experiences." *Dyslexia* 24 (3):207-19. doi: 10.1002/dys.1593.
- Kodan, H., and H. Akyol. 2018. "Effects of choral, repeated and assisted reading strategies on reading and reading comprehension skills of poor readers." *Egitim ve Bilim* 43 (193):159-79. doi: 10.15390/EB.2018.7385.
- Lawrence, J. F., A. M. Hagen, J. K. Hwang, G. Lin, and A. Lervåg. 2018. "Academic vocabulary and reading comprehension: exploring the relationships across measures of vocabulary knowledge." *Reading and Writing*:1-22. doi: 10.1007/s11145-018-9865-2.
- Lee, H. J., and M. Kim. 2015. "The correlation between metaphor comprehension and reading comprehension of poor readers in 4-6 grades." *Communication Sciences and Disorders* 20 (2):331-43. doi: 10.12963/csd.15238.
- Lopes, J. 2012. "Biologising reading problems: the specific case of dyslexia." *Contemporary Social Science: Journal of the Academy of Social Sciences* 7 (2):215-29. doi: 10.1080/21582041.2012.692098.
- Lopes, J., C. M. S. Gomes, C. R. Oliveira, and J. Elliott. 2016. "Selecting participants for dyslexia studies: Complexities and challenges " In *Dyslexia: Perspectives, Challenges and Treatment Options* edited by Marilyn Frazier, 145-58. Hauppauge, N.Y. USA: Nova Science Publishers.
- Luciano, M., A. J. Gow, A. Pattie, T. C. Bates, and I. J. Deary. 2018. "The Influence of Dyslexia Candidate Genes on Reading Skill in Old Age." *Behavior Genetics* 48 (5):351-60. doi: 10.1007/s10519-018-9913-3.
- Lyon, G. R., and L. C. Moats. 1997. "Critical Conceptual and Methodological Considerations in Reading Intervention Research." *Journal of Learning Disabilities* 30 (6):578-88. doi: 10.1177/002221949703000601.
- Machek, G. R., and J. M. Nelson. 2007. "How should reading disabilities be operationalized?: A survey of practicing school psychologists." *Learning Disabilities Research & Practice*:147-57. doi: 10.1111/j.1540-5826.2007.00239.x.

- McGuinness, D. 2005. *Language development and learning to read*. Cambridge: MIT Press.
- Moats, L. 2014. "What teachers don't know and why they aren't learning it: addressing the need for content and pedagogy in teacher education." *Australian Journal of Learning Difficulties* 19 (2):75-91. doi: 10.1080/19404158.2014.941093.
- Nergård-Nilssen, T., and K. Eklund. 2018. "Evaluation of the psychometric properties of "the Norwegian screening test for dyslexia". *Dyslexia* 24 (3):250-62. doi: 10.1002/dys.1577.
- Norton, E. S., S. D. Beach, and J. D. E. Gabrieli. 2015. "Neurobiology of dyslexia." *Current Opinion in Neurobiology* 30:73-8. doi: 10.1016/j.conb.2014.09.007.
- O'Donnell, P. S., and D. N. Miller. 2011. "Identifying students with specific learning disabilities: School psychologists' acceptability of the discrepancy model versus response to intervention." *Journal of Disability Policy Studies* (22):83-94. doi: 10.1177/1044207310395724.
- Orton, S. T. 1925. "Word-blindness' in school children." *Archives of Neurology and Psychiatry* 14:285-516. doi: doi:10.1001/archneurpsyc.1925.02200170002001.
- Patton, M. K. 2015. *Qualitative research & evaluation methods: Integrating theory and practice*. Thousand Oaks, CA: SAGE.
- Pennington, B. F. 2006. "From single to multiple deficit models of developmental disorders." *Cognition* 101 (2):385-413. doi: 10.1016/j.cognition.2006.04.008.
- Peters, L., and D. Ansari. 2019. "Are specific learning disorders truly specific, and are they disorders?" *Trends in Neuroscience and Education* 17. doi: 10.1016/j.tine.2019.100115.
- Peterson, R. L., and B. F. Pennington. 2015. "Developmental dyslexia." In *Annual Review of Clinical Psychology*, 283-307.
- Protopapas, A., and R. Parrila. 2018. "Is dyslexia a brain disorder?" *Brain Sciences* 8 (4). doi: 10.3390/brainsci8040061.
- Rice, M. , and G. Brooks. 2004. *Developmental dyslexia in adults: A research review*. London: NRDC.
- Richlan, F. 2014. "Functional neuroanatomy of developmental dyslexia: The role of orthographic depth." *Frontiers in Human Neuroscience* 8 (MAY). doi: 10.3389/fnhum.2014.00347.
- Ring, J., and J. L. Black. 2018. "The multiple deficit model of dyslexia: what does it mean for identification and intervention?" *Annals of Dyslexia* 68 (2):104-25. doi: 10.1007/s11881-018-0157-y.
- Rutter, M., and B. Maughan. 2005. "Dyslexia: 1965-2005." *Behavioural and Cognitive Psychotherapy* (33):389-402. doi: 10.1017/S1352465805002316.
- Ryder, D., and B. Norwich. 2018. "What's in a name? Perspectives of dyslexia assessors working with students in the UK higher education sector." *Dyslexia* 24 (2):109-27. doi: 10.1002/dys.1582.
- Schneider, W. J., and A. S. Kaufman. 2017. "Let's not do away with comprehensive cognitive assessments just yet." *Archives of Clinical Neuropsychology* 32 (1):8-20. doi: 10.1093/arclin/acw104.
- Scruggs, T. E., and M. A. Mastropieri. 2002. "On babies and bathwater: Addressing the problem of identification of learning disabilities." *Learning Disability Quarterly* (25):155-68. doi: 10.2307/1511299.
- Seidenberg, M. . 2017. *Language at the Speed of Sight, How We Read, Why So Many Can't, and What Can Be Done About It*. New York: Basic Books.
- Shaywitz, S. 2003. *Overcoming dyslexia: A new and complete science-based program for reading problems at any level*. New York: Vintage.

- Shaywitz, S., and B. A. Shaywitz. 2013. "Making a hidden disability visible: What has been learned from neurobiological studies of dyslexia." In *Handbook of learning disabilities* edited by H. L. Swanson, K. R. Harris and S. Graham, 643-57. New York: Guilford Press.
- Sjoblom, A. M., E. Eaton, and S. D. Stagg. 2016. "The effects of letter spacing and coloured overlays on reading speed and accuracy in adult dyslexia." *British Journal of Educational Psychology* 86 (4):630-9. doi: 10.1111/bjep.12127.
- Snowling, M. J., A. Lervåg, H. M. Nash, and C. Hulme. 2019. "Longitudinal relationships between speech perception, phonological skills and reading in children at high-risk of dyslexia." *Developmental Science* 22 (1). doi: 10.1111/desc.12723.
- Stagg, S. D., E. Eaton, and A. M. Sjoblom. 2018. "Self-efficacy in undergraduate students with dyslexia: a mixed methods investigation." *British Journal of Special Education* 45 (1):26-42. doi: 10.1111/1467-8578.12200.
- Stanovich, K. E. 1988. "Explaining the differences between the dyslexic and the garden variety poor reader: The phonological-core variable-difference model." *Journal of Learning Disabilities* 21:509-604.
- . 2005. "The future of a mistake: Will discrepancy measurement continue to make the learning disabilities field a pseudoscience?" *Learning Disability Quarterly* 28 (2):103-6. doi: 10.2307/1593604.
- Sternberg, R. J., and E. L. Grigorenko. 2002. "Difference scores in the identification of children with learning disabilities - It's Time to use a different method." *Journal of School Psychology* 40 (1):65-83. doi: 10.1016/S0022-4405(01)00094-2.
- Stuebing, K. K., A. E. Barth, P. J. Molfese, B. Weiss, and J. M. Fletcher. 2009. "IQ is not strongly related to response to reading instruction: a meta-analytic interpretation." *Exceptional Children* 76 (1):31-51.
- Su, M., J. Zhao, M. Thiebaut de Schotten, W. Zhou, G. Gong, F. Ramus, and H. Shu. 2018. "Alterations in white matter pathways underlying phonological and morphological processing in Chinese developmental dyslexia." *Developmental Cognitive Neuroscience* 31:11-9. doi: 10.1016/j.dcn.2018.04.002.
- Tilanus, E. A. T., E. Segers, and L. Verhoeven. 2019a. "Predicting responsiveness to a sustained reading and spelling intervention in children with dyslexia." *Dyslexia* 25 (2):190-206. doi: 10.1002/dys.1614.
- . 2019b. "Responsiveness to Intervention after Second versus Third Grade Diagnosis of Dyslexia." *Reading and Writing Quarterly*. doi: 10.1080/10573569.2019.1667929.
- Tong, X., C. McBride, J. C. M. Lo, and H. Shu. 2017. "A Three-Year Longitudinal Study of Reading and Spelling Difficulty in Chinese Developmental Dyslexia: The Matter of Morphological Awareness." *Dyslexia* 23 (4):372-86. doi: 10.1002/dys.1564.
- Tunmer, W., and K. Greaney. 2010. "Defining dyslexia." *Journal of Learning Disabilities* 43 (3):229-43. doi: 10.1177/0022219409345009.
- Uno, A., T. N. Wydell, N. Haruhara, M. Kaneko, and N. Shinya. 2009. "Relationship between reading/writing skills and cognitive abilities among Japanese primary-school children: Normal readers versus poor readers (dyslexics)." *Reading and Writing* 22 (7):755-89. doi: 10.1007/s11145-008-9128-8.
- Vellutino, F. R., J. M. Fletcher, M. J. Snowling, and D. M. Scanlon. 2004. "Specific reading disability (dyslexia): what have we learned in the past four decades?" *Journal of Child Psychology and Psychiatry* 45 (1):2-40.

- Ward-Lonergan, J. M., and J. K. Duthie. 2018. "The state of dyslexia: Recent legislation and guidelines for serving school-age children and adolescents with dyslexia." *Language, Speech, and Hearing Services in Schools* 49 (4):810-6. doi: 10.1044/2018_LSHSS-DYSLC-18-0002.
- Williams, J. L., J. Miciak, L. McFarland, and J. Wexler. 2016. "Learning Disability Identification Criteria and Reporting in Empirical Research: A Review of 2001–2013." *Learning Disabilities Research and Practice* 31 (4):221-9. doi: 10.1111/ldrp.12119.
- Xia, Z., F. Hoeft, L. Zhang, and H. Shu. 2016. "Neuroanatomical anomalies of dyslexia: Disambiguating the effects of disorder, performance, and maturation." *Neuropsychologia* 81:68-78. doi: 10.1016/j.neuropsychologia.2015.12.003.